Department of Mathematics

Predicting How Many Points a Famous NBA Player Will Score in a Game

Crystal Brock
Advisor: Dr. Maher Qumsiyeh

Abstract

Stephen(Steph) Curry is a player in the National Basketball Association (NBA), and plays as a point guard for the Golden State Warriors. The purpose of this project is to predict, based on a number of factors (independent variables) from the 2015-2016 season, the amount of points Steph Curry will score in any given game. The independent variables (regressors) chosen that could have an effect on the number of points scored are minutes played in a specific game, field goal percentage, three point percentage, free throw percentage, rebounds, assists, blocks, steals, fouls, and turnovers. The data was obtained from the ESPN website. The statistical software package, SPSS, was used to analyze the data and obtain a good prediction model.

Model development

We want to predict the number of points scored in a game by Steph Curry (dependent variable (y)) by using several regressors (independent variables) such as field goal percentage, minutes played in a game, three point percentage, free throw percentage, rebounds, assists, blocks, steals, fouls and turnovers.

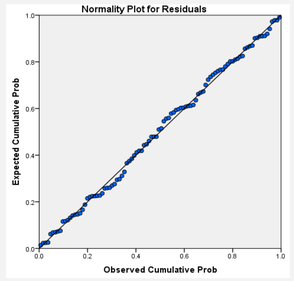
Coefficients[a]

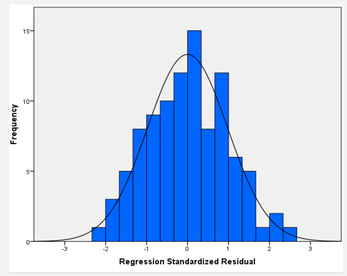
Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2.556	.314		8.150	.000
	Field Goal Percent	5.622	.625	.676	8.990	.000
2	(Constant)	-1.240	.437		-2.836	.006
	Field Goal Percent	6.474	.446	.779	14.516	.000
	Minutes	.099	.010	.539	10.040	.000
3	(Constant)	-1.092	.422		-2.590	.011
	Field Goal Percent	5.074	.626	.610	8.110	.000
	Minutes	.096	.009	.524	10.151	.000
	3 point percent	1.443	.471	.227	3.064	.003
4	(Constant)	-1.010	.417		-2.422	.017
	Field Goal Percent	4.894	.622	.589	7.866	.000
	Minutes	.094	.009	.514	10.058	.000
	3 point percent	1.424	.464	.224	3.070	.003
	FT percent	.059	.029	.103	2.016	.047

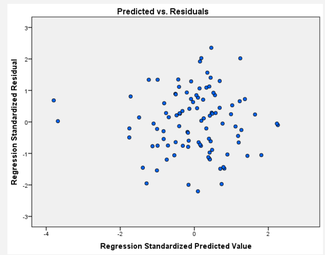
a. Dependent Variable: sqrtpoints

Using the stepwise procedure in the statistical software, SPSS, we are able to develop the best model and SPSS gives the above output.

By using the Box-Cox method to obtain the best transformation of y, I was able to determine that the was best.







Model Adequacy Checking

The Normality Plot for Residuals shows that the data collected lies approximately along the straight line, and the histogram depicted also shows that the residuals are approximately normally distributed.

The Predicted vs Residuals plot is also randomly scattered which is ideal after the transformation of y into $\sqrt{y}$ was performed.

The table below gives the F-statistic along with the p-value that proves the model exists and is valid.

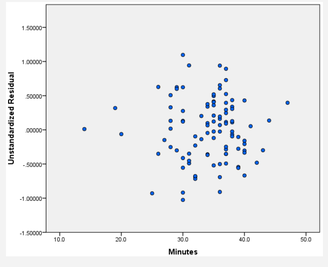
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	67.686	4	16.921	78.070	.000[b]
	Residual	20.157	93	.217		
	Total	87.843	97			

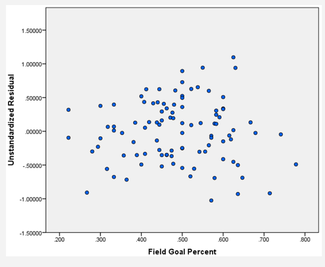
a. Dependent Variable: sqrtpoints

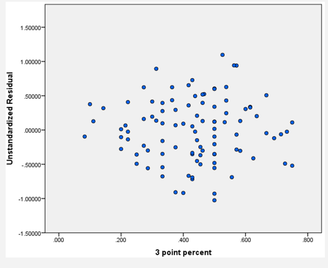
b. Predictors: (Constant), FT percent, Minutes, 3 point percent, Field Goal Percent

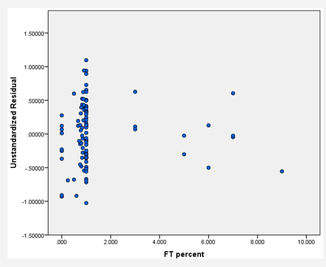
When plotting each regressor against the residuals; minutes, field goal percentage, and three point percentage all showed the same random scatter that is ideal.

However, free throw percentage did not show the ideal scatte r and the observed values are all clustered around a few percentages and that is due to the variable having a few common values.









Final Model

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-1.010	.417		-2.422	.017
	Minutes	.094	.009	.514	10.058	.000
	Field Goal Percent	4.894	.622	.589	7.866	.000
	3 point percent	1.424	.464	.224	3.070	.003
	FT percent	.059	.029	.103	2.016	.047

a. Dependent Variable: sqrtpoints

Equation (1) : $= -1.01 + 0.094x_1 + 4.894x_2 + 1.424x_3 + 0.059x_4$

Using the above equation developed from the significant variables and their coefficients in the above table, we are able to determine that the model will explain about 76.1% of the variations in .

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.878[a]	.771	.761	.46556

a. Predictors: (Constant), FT percent, Minutes, 3 point percent, Field Goal Percent

b. Dependent Variable: sqrtpoints

Conclusion

We were able to develop a model given by equation (1) that can be used to predict the square root of points for a given game relying on the factors of minutes, field goal percentage, three point percentage, and free throw percentage. In order to obtain the predicted number of points, just square the predicted value given by the model.